Encoding Mixtepec-Mixtec Etymology in TEI

Jack Bowers

jack.bowers@oeaw.ac.at

Austrian Center for Digital Humanities (ACDH)
Inria - Team ALMAnaCH, Paris
École Pratique des Hauts Études, Paris

Laurent Romary

Inria - Team ALMAnaCH, Paris
Berlin Brandenburgische Akademie der Wissenschaften (BBAW)

Inria
INVENTEURS DU MONDE NUMÉRIQUE

ÖAW
AUSTRIAN ACADEMY OF SCIENCES

Mixtepec-Mixtec

- Sa'an Savi "*rain language*"
- ISO 639-3 code: 'mix'
- San Juan de Mixtepec - Juxtlahuaca district (Oaxaca, MEX)
- "Vigorous" status but highly under-resourced
- Oto-Manguean, Mixtecan, Mixtec-Cuicatec, Mixtepec-Mixtec
- Tonal
- Spoken data mostly collected in sessions working with speakers from a small village called Yucunani in the San Juan Mixtepec municipality
- Estimated (+-9,000 -10,000 speakers)
  Source: (INEGI, 2010)
- Phonology has been studied by Pike and Ibach (1978); Paster and Beam de Azcona (2004-2007);
- Beckman and Nieves-SIL (2005-current) published booklets and are working on developing orthography

Introduction to Project

- Language documentation: (*PhD project)*
- Primary sources of language data:
  - speaker consultations (recordings, new written material..)
  - +- 40 Children's Booklets (SIL)
  - Public sources (You Tube, *other*)
  - examples from academic papers
- Goals:
  - TEI Corpus
  - Linguistic descriptions
  - TEI Dictionary *(actually 2 dictionaries, 1 general, 1 inflectional)*
  - *(Etymology)* would like to create data contents and structure that can be copied and integrated into treament of related languages

Overview of Issues & Sources in Mixtec Etymology

- Sense based changes (*metaphor, metonomy, grammaticalization*) identified in related Mixtec languages important for cognitive linguistics (Brugman, 1983; Beckman, 1995;…)
  - *especially body-part terms*

Sources

- Reconstructed Proto-Mixtecan from Longacker
- Loanwords easily identifiable (mostly from Spanish)
  - *phonological changes evident in these*
- Oldest sources of documentation of Mixtec languages from 1592

Issues, Challenges

- *Need to balance personal/academic/intelectual interests with the practicalities and needs of the language community*
- *Need Mixtec editors to provide Mixtec language versions of content (requires new use of language vocabulary!)*
- Other issues: *one man job*

Overview of Components of Etymological Markup in TEI

Bowers & Romary (2016) propose expansion and refinement of etymology section of the TEI dictionary

TEI Lex0 Etym (Bowers & al., 2018) builds off this…..

- Etymology element (<etym>): structuring etymology processes through typing and recursivity
- Typology of etymological processes
- Etymons and their forms
- Related forms (cognates, and others)
- Temporality of etymological processes
- Bibliographical references in etymologies
- Prose description of etymological process and content

Etymological Features: Cognates

Related forms extracted (manually) from academic papers on other Mixtec varieties

....

(Chalcatongo Mixtec: San Miguel El Grande) ***šinì*** *(Macaulay, 1996);*

(Ayutla Mixtec) ***shīhih*** *(Hills, 1990)*;

(Coatzospan Mixtec) ***ʃɪɲī*** *(Padget, 2017)*;

(Guadalupe Nundaca) ***ʃɪɲī*** *(Macaulay, 1996)*;

```
<cit type="cognate">
  <lang>Chalcatongo Mixtec</lang>
  <usg type="geo">
    <placeName>San Miguel El Grande</placeName>
  </usg>
  <form>
    <pron notation="trans-macaulay-mig" xml:lang="mig">šinì</pron>
  </form>
  <ref type="source" target="#Macaulay-ChalcatongoMixtec-1996">(Macaulay, 1996)</ref>
</cit>

<cit type="cognate">
  <lang>Ayutla Mixtec</lang>
  <form>
    <pron notation="trans-hill-1990-miy" xml:lang="miy">shīhih</pron>
  </form>
  <ref target="#Hills-AyutlaMixtec-1990">(Hills, 1990)</ref>
</cit>

<cit type="cognate">
  <lang>San Martín Duraznos</lang>
  <form>
    <pron notation="ipa" xml:lang="smd">ʃɪɲī</pron>
  </form>
  <ref target="#Padgett-2017">(Padget, 2017)</ref>
</cit>

<cit type="cognate">
  <lang>Guadalupe Nundaca</lang>
  <form>
    <pron notation="ipa" xml:lang="gna">ʃɪɲī</pron>
  </form>
  <ref target="#Padgett-2017">(Padget, 2017)</ref>
</cit>
```

Etymological Features: Cognates & Bibliographic Sources

<bibl xml:id="Macaulay-ChalcatongoMixtec-1996">
Macaulay, M. (1996). A grammar of Chalcatongo Mixtec (University of California publications in linguistics ed.). University of California Press.**</bibl>**

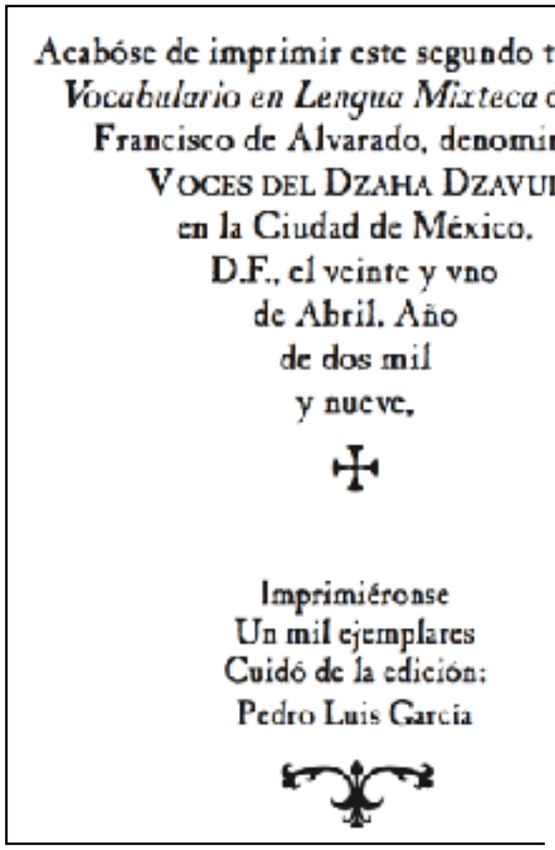
<bibl xml:id="Hills-AyutlaMixtec-1990">
Hills, Robert A. 1990. A syntactic sketch of Ayutla Mixtec. Studies in the Syntax of Mix-tecan Languages, vol. 2, Summer Institute of Linguistics and the University of Texas at Arlington Publications in Linguistics, no. 90, ed. C. Henry Bradley and Barbara E. Hollenbach, pp. 1-260. Dallas: Summer Institute of Linguistics and the University of Texas at Arlington. **</bibl>**

<bibl type="thesis" xml:id="Padgett2017">
Padgett, E. (2017). Tools for assessing relatedness in understudied language varieties: a survey of Mixtec varieties in Western Oaxaca, Mexico. University of North Dakota. **</bibl>**

```
<cit type="cognate">
  <lang>Chalcatongo Mixtec</lang>
  <usg type="geo">
    <placeName>San Miguel El Grande</placeName>
  </usg>
  <form>
    <pron notation="trans-macaulay-mig" xml:lang="mig">šinì</pron>
  </form>
  <ref type="bibl" target="#Macaulay-ChalcatongoMixtec-1996">
      (Macaulay, 1996)</ref>
</cit>
<cit type="cognate">
  <lang>Ayutla Mixtec</lang>
  <form>
    <pron notation="trans-hill-1990-miy" xml:lang="miy">shīhih</pron>
  </form>
  <ref type="bibl" target="#Hills-AyutlaMixtec-1990">(Hills, 1990)</ref
</cit>
<cit type="cognate">
  <lang>San Martín Duraznos</lang>
  <form>
    <pron notation="ipa" xml:lang="smd">ʃīɲī</pron>
  </form>
   <ref type="bibl" target="#Padgett-2017">(Padget, 2017)</ref>
</cit>
<cit type="cognate">
  <lang>Guadalupe Nundaca</lang>
  <form>
    <pron notation="ipa" xml:lang="gna">ʃīɲī</pron>
  </form>
  <ref type="bibl" target="#Padgett-2017">(Padget, 2017)</ref>
</cit>
```

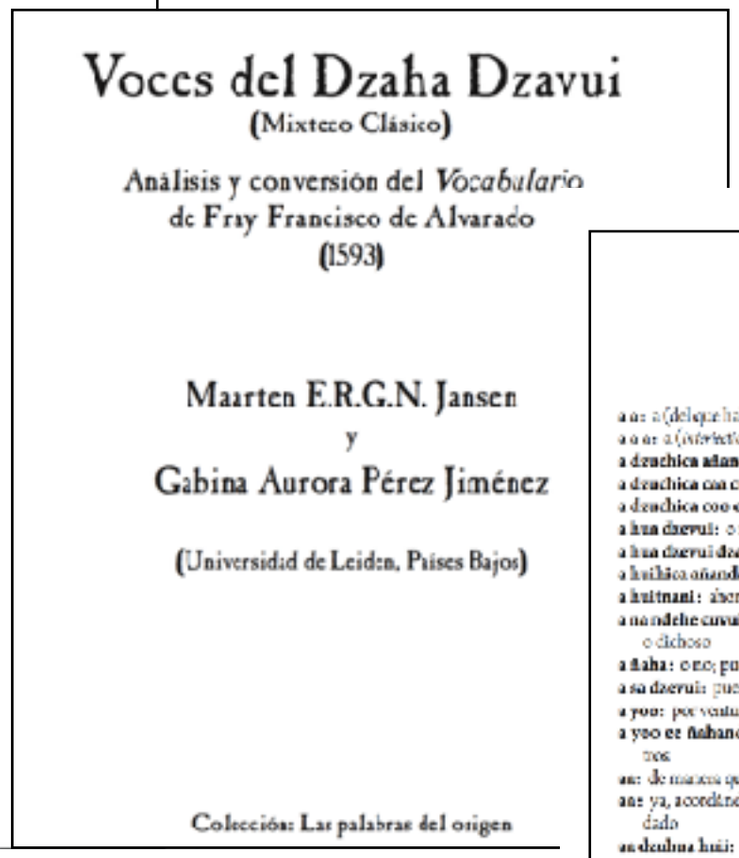
Integrating Sources: Colonial Mixtec Dictionary

- *GROBID dictionaries (Khemakhem et al., 2017)*



Acabóse de imprimir este segundo tomo del
Vocabulario en Lengua Mixteca de Fray
Francisco de Alvarado, denominado
VOCES DEL DZAHA DZAVUI.
en la Ciudad de México,
D.F., el veinte y vno
de Abril. Año
de dos mil
y nueve.

Imprimiéronse
Un mil ejemplares
Cuidó de la edición:
Pedro Luis García



Voces del Dzaha Dzavui
(Mixteco Clásico)

Análisis y conversión del *Vocabulario*
de Fray Francisco de Alvarado
(1593)

Maarten E.R.G.N. Jansen
y
Gabina Aurora Pérez Jiménez

(Universidad de Leiden, Países Bajos)

Colección: Las palabras del origen

A

a ao: a (de que habla otro en malefícia)
a o ao: a (*interfecto adversativi*)
a deuchica adandaa: poco más o menos
a deuchica con cuvui: poco más o menos
a deuchica coo cuvui: poco más o menos
a hua dzevui: o no
a hua dzevui dzevua: pues no
a huihica añandoo: poco más o menos
a huitnani: aben poco
a nandehe cuvui ndatu aicay: o bienaventurado, o dichoso
a ñaha: o no; pues no
a sa dzevui: pues no
a yoo: por ventura alguno
a yoo ee ñahando: por ventura alguno de vosotros
aa: de manera que
aa: ya, acordándoseme lo que se me había olvidado
aa deuhua huii: así, así (acordándose)
aa ndica huii: ya (acordándoseme lo que se me había olvidado)
adai: o (*dyuntiv*); por ventura; quizás
adai: suave cosa
adai cuvui: o (*dyuntiv*); por ventura; quizás
adai q cuvui: por ventura
adai yoo: por ventura alguno
adai yoo ee ñahando: por ventura alguno de vosotros
ahua: ay, quejándose la mujer
ama: así
ama: bienevi (otorgando); sí
amana: ¿cuándo? (*adverbii interrogativo*) ¿en qué tiempo?
amana cuiya: en algún tiempo
amana cuvui huatu iaiando: cuandoquiera que quisieres
amana na ndita ñamana nuundo: ¿cuándo has de despertar?
amana na ndotondo: ¿cuándo has de despertar?
amana na tahui iñando: ¿cuándo has de volver en tí?; ¿cuándo has de despertar?
amana quevui: en algún tiempo
amana quevui sa cuvui iaindo: cuandoquiera que quisieres
amanaca: ¿cuándo, en qué tiempo?
amani: de tarde en tarde o rara vez
andaya: infierno, lugar de dañados
andevui: cielo
andevui isi ndaa tiñoo: cielo estrellado
angel nicoo coo ndaa ndita ña: ángel de mi guarda
angel yondoca ñaha: ángel de mi guarda
aniñe: palacio
anuhu: abismo; centro de la tierra: infierno, lugar de dañados
anuhu: profundo
anuhu naa: profundo
anuhu naa: infierno, lugar de dañados
anuhu ndahui: infierno, lugar de dañados
ataa: ojalá
ata: amarga cosa; áspero al gusto
aya: amarga cosa; áspero al gusto

Provides new source of etymons!!

Jansen, M. E. R. G. N., & Perez, G. A. (2009). Voces del Dzaha Dzavui (mixteco clásico). Análisis y Conversión del Vocabulario de fray Francisco de Alvarado (1593).

Integrating Sources: Colonial Mixtec Dictionary

andevui: cielo

```
<entry xml:id="andevui">
  <form type="lemma">
    <orth xml:lang="nds-x-clmx">andevui</orth>
  </form>
  <pc>:</pc>
  <sense>
    <def xml:lang="es">cielo</def>
  </sense>
</entry>
```

```
<entry xml:id="sky">
  <form type="lemma">
    <orth xml:lang="mix">antivi</orth>
    <gramGrp>
      <pos>noun</pos>
    </gramGrp>
  </form>
  <sense corresp="http://dbpedia.org/resource/Sky">
    <usg type="domain">Meteorology</usg>
    <cit type="translation">
      <form>
        <orth xml:lang="en">sky</orth>
      </form>
    </cit>
    <cit type="translation">
      <form>
        <orth xml:lang="es">cielo</orth>
      </form>
    </cit>
  </sense>
  <etym type="inheritance">
    <cit type="etymon">
      <form>
        <orth xml:lang="nds-x-clmx">andevui</orth>
      </form>
      <gloss xml:lang="es">cielo</gloss>
      <ref type="bibl" target="#VOCESvocab-tei.xml">
        Francisco de Alvarado</ref>
    </cit>
  </etym>
</entry>
```

Jansen, M. E. R. G. N., & Perez, G. A. (2009). Voces del Dzaha Dzavui (mixteco clásico). Análisis y Conversión del Vocabulario de fray Francisco de Alvarado (1593).

Etymological Processes: Inheritance

Where existing in a past source or where we reconstruct a form based on comparisson of cognates, we may place a Proto-Mixtecan form in the <etym> section as the etymon.

```
<entry xml:id="ntuchi">
  <form type="lemma">
    <orth xml:lang="mix">ntuchi</orth>
    <pron xml:lang="mix" notation="ipa">ndùʧí</pron>
    <gramGrp>
      <pos>noun</pos>
    </gramGrp>
  </form>

  ....
  <etym type="inheritance">
    <cit type="etymon" cert="medium">
      <form>
        <pron notation="ipa" orig="*ⁿdu³ⁿdi⁴" xml:lang="nds-PMx">*ⁿdu˩ⁿdi˥</pron>
      </form>
      <ref source="#Longacre1961">(Longacre and Millon, 1961)</ref>
    </cit>
  </etym>

</entry>
```

BEAN (?) [17]. M-SM, SE ^{n}dūčī. J ^{n}duti. C nūūnē/nūūnī. (Also means kidney in M-SM, SE and J.) T-Ch ru^{3}ne$^{4-5}$ (ga^{4}ʔnï ʔi$^{4-3}$) large black beans, ru^{3}ne$^{4-5}$ (gi^{4}ci^{3}) small vari-colored beans, zi^{3}-ru^{4}ne^{4} kidneys. A ntæ, ntæ ʔ kidneys. PMx *^{n}du^{3n}di^{4}. PAMx *^{n}di. This is a strong Proto-Mixtecan and Proto-Amuzgo-Mixtecan set. Cf. comment on

Etymological Processes: Borrowing

There is a significant amount of loanwords in the language (the vast majority of which are from Spanish).

```
<entry xml:id="mother">
  <form type="lemma">
    <orth xml:lang="mix">maa</orth>
    <pron xml:lang="mix" notation="ipa">māā</pron>
    <gramGrp>
      <pos>noun</pos>
    </gramGrp>
  </form>
  <sense>
   <!— other info here -->
    <cit type="translation">
      <form>
        <orth xml:lang="en">mother</orth>
      </form>
    </cit>
  </sense>
   <etym type="borrowing">
     <cit type=“etymon">
      <lang xml:lang="en">Sp.</lang>
       <form>
         <orth xml:lang=“es">madre</orth>
         <pron xml:lang="es">'madre</pron>
       </form>
     </cit>
   </etym>
</entry>
```

Etymological Processes: Compounding

```
<entry xml:id="num-16" type="compound">
  <form type="lemma">
    <orth xml:lang="mix">tsa'un in</orth>
      ....
  </form>
    <gramGrp>
      <pos>cardinalNumber</pos>
    </gramGrp>
  <sense>
    <usg type="domain">CardinalNumber</usg>
    <cit type="translation">
      <form>
        <orth xml:lang="en">sixteen</orth>
      </form>
    </cit>
    <cit type="translation">
      <form>
        <orth xml:lang="es">dieciséis</orth>
      </form>
    </cit>
  </sense>
  <etym>
    .....
  </etym>
</entry>
```

```
<etym type="compounding">

  <seg type="desc" xml:lang="en">From:</seg>
  <seg type="desc" xml:lang="es">De:</seg>
  <seg type="desc" xml:lang="mix">Ña:</seg>

  <cit type="etymon" corresp="#num-15">
    <form>
      <orth xml:lang="mix">tsa'un</orth>
    </form>
    <gloss xml:lang="en">fifteen</gloss>
    <gloss xml:lang="es">quinze</gloss>
  </cit>

  <pc>+</pc>

  <cit type="etymon" corresp="#num-1">
    <form>
      <orth xml:lang="mix">in</orth>
    </form>
    <gloss xml:lang="en">one</gloss>
    <gloss xml:lang="es">uno</gloss>
  </cit>

</etym>
```

Etymological Processes: Compounding

```
<etym type=“compounding">

  <seg type="desc" xml:lang=“en">From:</seg>
  <seg type="desc" xml:lang="es">De:</seg>
  <seg type="desc" xml:lang=“mix">Ña:</seg>

  <cit type="etymon" corresp="#num-15">
    <form>
      <orth xml:lang="mix">tsa'un</orth>
    </form>
    <gloss xml:lang="en">fifteen</gloss>
    <gloss xml:lang="es">quinze</gloss>
  </cit>

  <pc>+</pc>

  <cit type=“etymon" corresp="#num-1">
    <form>
      <orth xml:lang="mix">in</orth>
    </form>
    <gloss xml:lang="en">one</gloss>
    <gloss xml:lang="es">uno</gloss>
  </cit>

</etym>
```

User oriented view output

Mixtec version

tsa’un in *(num.)*
(Es.) dieciséis; (En.) fifteen
Ña: *tsa'un* + *in*

Spanish version

tsa’un in *(num.)*
(Es.) dieciséis; (En.) fifteen
De: *tsa'un (quinze)* + *in* (uno)

English version

tsa’un in *(num.)*
(Es.) dieciséis; (En.) fifteen
From: *tsa'un (fifteen)* + *in* (*one*)

Etymological Processes: Derivation

```
<entry xml:id="sharpen">
  <form type="lemma">
    <orth xml:lang="mix">ntasaxeen</orth>
  </form>
  <gramGrp>
    <pos>verb</pos>
    <gram type="transitivity">trans</gram>
    <gram>causative</gram>
    <gram>iterative</gram>
  </gramGrp>
  <sense>
    <cit type="translation">
      <form>
        <orth xml:lang="en">sharpen</orth>
      </form>
    </cit>
    <cit type="translation">
      <form>
        <orth xml:lang="es">afilar</orth>
      </form>
    </cit>
  </sense>
  <etym type="derivation">
  ……
  </etym>
</entry>
```

nta-

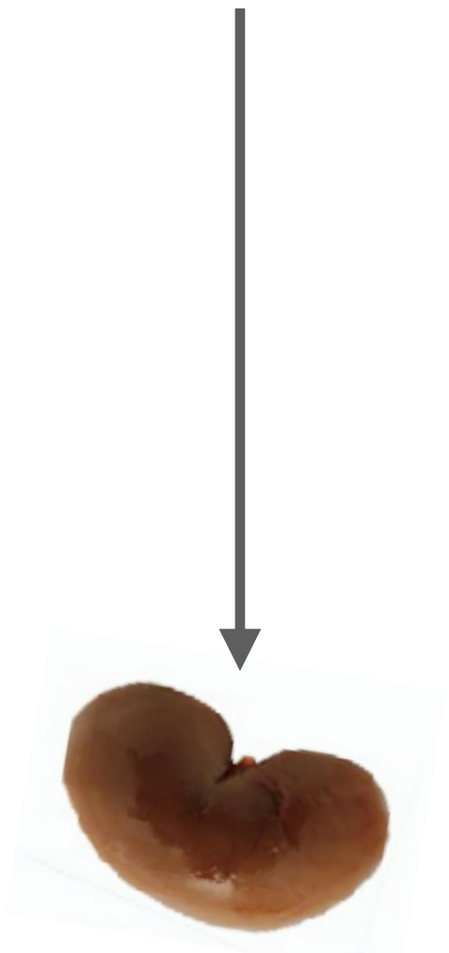
sa-

xeen

```
<etym type="derivation">
  ….
  <cit type="etymon">
    <form>
      <orth xml:lang="mix">nta-</orth>
    </form>
    <gramGrp>
      <gram>prefix</gram>
       <gram>iterative</gram>
    </gramGrp>
  </cit>
  <pc>+</pc>
  <cit type="etymon">
    <form>
      <orth xml:lang="mix">sa-</orth>
    </form>
    <gramGrp>
      <gram>prefix</gram>
       <gram>causative</gram>
    </gramGrp>
  </cit>
  <pc>+</pc>
  <cit type="etymon">
    <form>
      <orth xml:lang="mix">xeen</orth>
    </form>
    <gramGrp>
      <pos>adj</pos>
    </gramGrp>
    <gloss xml:lang="en">dangerous</gloss>
    <gloss xml:lang="es">peligroso</gloss>
  </cit>
</etym>
```

Etymological Processes: Metaphor

source concept



target concept

```
<sense xml:id="bean" corresp="http://dbpedia.org/resource/Pinto_bean" n="1">
    <usg type="domain">Legume</usg>
    <usg type="domain">stapleFoods</usg>
    <cit type="translation">
      <form>
        <orth xml:lang="en">bean</orth>
      </form>
    </cit>
    <cit type="translation">
      <form>
        <orth xml:lang="es">frijol</orth>
      </form>
    </cit>
  </sense>
  <sense xml:id="kidney" corresp="http://dbpedia.org/resource/Kidney" n="2">
    <usg type="domain">InternalOrgan</usg>
    <cit type="translation">
      <form>
        <orth xml:lang="en">kidney</orth>
      </form>
    </cit>
    <cit type="translation">
      <form>
        <orth xml:lang="es">riñón</orth>
      </form>
    </cit>
    <etym type="metaphor">
     ......
    </etym>
  </sense>
```

Etymological Processes: Metaphor

```
<sense xml:id="bean" corresp="http://dbpedia.org/resource/Pinto_bean" n="1">
    <usg type="domain">Legume</usg>
    <usg type="domain">Food</usg>
     …..
  </sense>
  <sense xml:id="kidney" corresp="http://dbpedia.org/resource/Kidney" n="2">
  <!-- translations here -->
      <etym type="metaphor">
        <seg type="desc" resp="#JB" xml:lang="en">The motivating source of the metaphor is
the similarity in shape and color between the source concept 'bean' and the target object 'kidney'. The
directionality of 'bean' → 'kidney' is assumed based on the fact that the later is a basic food staple of
Mixtec people and that according to the theory of embodiment, we would expect polysemy to proliferate
in the direction of: more basic/prominent → less basic/prominent. While body parts are amongst the
most prominent in creating polysemies and new terms, the kidney is an internal organ and thus it is likely
a less prominent concept in the embodied experience of the early Mixtecan speakers whose usage gave
rise to this lexicalization.</seg>
            <cit type="etymon" corresp="#bean">
             <sense>
               <usg type="domain">Food</usg>
               <usg type="domain">InternalOrgan</usg>
               <gloss xml:lang="en">bean</gloss>
               <gloss xml:lang="es">frijol</gloss>
             </sense>
            </cit>
          </etym>
        </sense>
```

Note: hope to further systematize the aspects of: active source and target salient concepts (ontology needed)

Etymological Processes: Metonymy

kiti *(n.)*
1) (Es.) animal; (En.) animal
2) (Es.) caballo; (En.) horse
……

ANIMAL is a hypernymOf HORSE

(implicit in data)

HORSE is a hypoymOf ANIMAL

```
<entry xml:id="animal-horse">
  <form type="lemma">
    <orth xml:lang="mix">kiti</orth>
    <pron xml:lang="mix" notation="ipa">kìt̪í</pron>
    <gramGrp>
      <pos>noun</pos>
    </gramGrp>
  </form>
  <sense xml:id=“animal" corresp="http://dbpedia.org/resource/Animal" n="1">
    <usg type="domain">livingBeing</usg>
    <cit type="translation">
      <form><orth xml:lang="en">animal</orth></form>
    </cit>
    …..
  </sense>
```

```
  <sense xml:id=“horse”  corresp="http://dbpedia.org/resource/Horse" n="2">
    <usg type=“domain">Animal</usg>
    <xr type=“hyponymOf">
      <ref xml:lang=“mix">kiti</ref>
      <ref xml:lang=“en">animal</ref>
    </xr>
      <cit type="translation">
       <form><orth xml:lang="en">horse</orth></form>
      </cit>
      <cit type="translation" xml:lang="es">
        <form><orth xml:lang="es">caballo</orth></form>
      </cit>
      <etym type="metonymy" subtype=“categoryForMember">
       ….
      </etym>
    </sense>
</entry>
```

Etymological Processes: Metonymy

kiti *(n.)*
1) (Es.) animal; (En.) animal
2) (Es.) caballo; (En.) horse
……

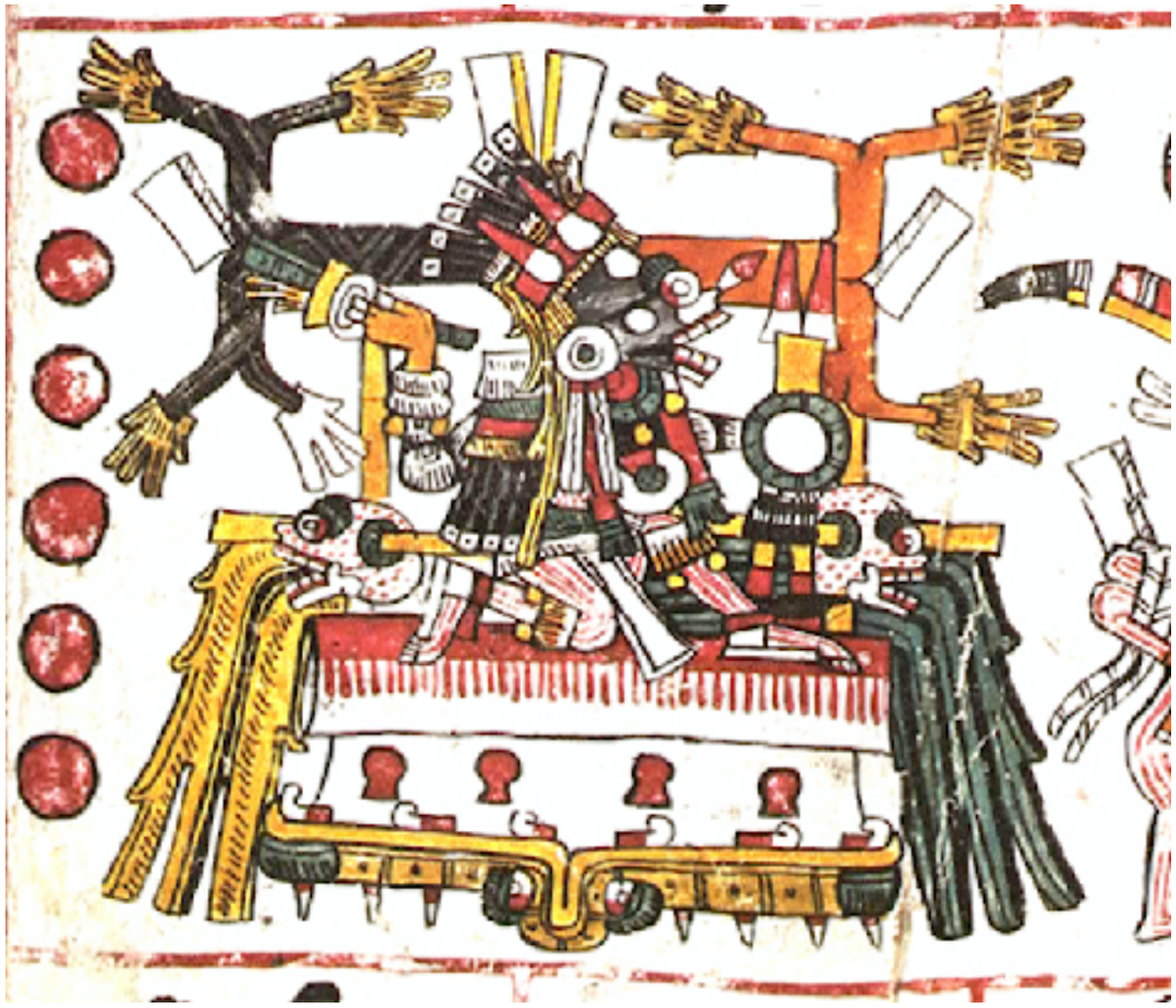
```
<sense xml:id="horse" corresp="http://dbpedia.org/resource/Horse" n="2">
  <usg type="domain">Animal</usg>
  <cit type="translation">
    <form><orth xml:lang="en">horse</orth></form>
  </cit>
  <cit type="translation" xml:lang="es">
    <form><orth xml:lang="es">caballo</orth></form>
  </cit>

  <xr type="hyponymOf">
    <ref xml:lang="mix">kiti</ref>
    <ref xml:lang="en">animal</ref>
  </xr>

  <etym type="metonymy" subtype="categoryForMember">
    <seg type="desc" resp="#JB" xml:lang="en">
      In this lexical item, the language reflects the history, since there were no
      horses in Mexico until the arrival of the Spanish (sometime after
      <date notBefore="1521">1521</date>), there was naturally no Mixtecan
      word for 'horse', thus the categorical noun meaning 'animal' was used to
      describe the unnamed animal.</seg>

    <cit type="etymon" corresp="#animal">
      <sense>
        <gloss xml:lang="en">animal</gloss>
        <gloss xml:lang="es">animal</gloss>
      </sense>
    </cit>
  </etym>
</sense>
```

Etymological Processes: Multiple Processes



Mixtec Codex Borgia

nta'a yutu
hand/arm tree
'tree branch'
'rama'

nta’a yutu
hand/arm tree
‘tree branch’
‘rama’

```
<entry xml:id="tree-branch" type="compound">
  <form type="lemma">
    <orth>nta'a yutu</orth>
    <pron notation="ipa">ndà?á jùtú</pron>
    <gramGrp>
      <gram>noun</gram>
    </gramGrp>
  </form>
  <sense>
    <graphic url="../imgs/apoala_tree.jpg"/>
    <usg type="domain">Botany</usg>
    <xr type="meronymOf">
      <ref xml:lang="mix">yutu</ref>
      <ref xml:lang="en">tree</ref>
    </xr>
    <cit type="translation">
      <form><orth xml:lang="en">branch</orth></form>
    </cit>
    <cit type="translation">
      <form><orth xml:lang="es">rana</orth></form>
    </cit>
  </sense>
  <etym type=“compounding">
    ……..
  </etym>
</entry>
```

```
<etym type="compounding">
  <etym type="metaphor">
    <cit type="etymon">
      <form>
        <orth>nta'a</orth>
      </form>
      <sense>
        <usg type="domain">HumanAntomy</usg>
        <xr type="meronymOf">
          <ref xml:lang="mix">kuñu</ref>
          <ref xml:lang="en">body</ref>
        </xr>
        <gloss xml:lang="en">hand</gloss>
        <gloss xml:lang="es">mano</gloss>
      </sense>
    </cit>
  <pc>+</pc>
  </etym>
  <cit type="etymon">
    <form>
      <orth>jutu</orth>
    </form>
    <gloss xml:lang="en">tree</gloss>
    <gloss xml:lang="es">arbol</gloss>
  </cit>
</etym>
```

Conclusion

- Creating etymological resources in TEI for an indigenous language is challenging given limited sources, the need to balance the perspectives and target audiences, e.g. linguistic (academic), technological, and especially language community

- Encoding etymology for a Mixtec language required and inspired the expansion of the features of TEI dealing with etymology

- Need to further be able to systematically extract and integrate cognate and (very limited) historical and reconstructed Proto-Mixtecan data

- Doing the above in TEI can hopefully provide an extensible set of data for other project in related languages to make use of and add to the resources created in this project